

## Makalah Seminar Tugas Akhir

### APLIKASI PENGENALAN UCAPAN DENGAN JARINGAN SYARAF TIRUAN

#### PROPAGASI BALIK UNTUK PENGENDALIAN *SMART WHEELCHAIR*

Dewanto Arby Hudhaya <sup>[1]</sup>, Achmad Hidayatno, ST, MT <sup>[2]</sup>, Iwan Setiawan, ST, MT <sup>[2]</sup>

Jurusan Teknik Elektro, Fakultas Teknik, Universitas Diponegoro

Jln. Prof. Sudharto, Tembalang, Semarang, Indonesia

#### ABSTRAK

*Speech recognition process can be done in many ways one of them with artificial neural networks. In order to be easily understood and to understand, it would require some method of characteristics extraction methods such as by LPC and Fourier transformation. Linear Predictive Coding is one tool in signal processing or signal analysis to obtain the unique characteristics of each sound pattern. While the Fourier transform is used to clarify the characteristics of each pattern as it can provide information that is presented in the frequency domain of both discrete and continuous.*

*ANN (Artificial Neural Networks) are widely used for various applications of pattern recognition. The ability of learning from training data and generalize to the situation / condition is new is the fundamental reason why ANN is used. In this Final Project, ANN used is the Back Propagation.*

*Furthermore, speech recognition results are translated into ASCII characters are then sent to a wheelchair through a serial connection that was identified as five voice commands. The voice commands consist of commands forward, backward, right, left, stop.*

*Test results show that the recognition of new data lower then the exercise data. From some variation of test, obtained the best network tes\_fix3, with speech recognition percentage of respondents in the database is 87%, and speech recognition percentage of respondents outside the database is 85,33%.*

**Key Word :** *speech recognition, LPC, Fourier, ANN, Back Propagation, serial, wheelchair*

#### I. PENDAHULUAN

##### 1.1 Latar Belakang

Pada era teknologi sekarang ini, teknik kendali banyak mengalami perkembangan. Selain teknik kendali, teknologi telekomunikasi juga mengalami perkembangan yang pesat pula. Dengan demikian, aplikasi penggabungan kedua bidang ini juga semakin meningkat. Salah satunya adalah *speech recognition* untuk bidang pengendalian. Banyak sekali aplikasi yang diterapkan seperti pengendalian robot dengan menggunakan ucapan manusia. Hal ini bertujuan untuk mempermudah pekerjaan manusia dalam berbagai bidang. Misalkan seseorang yang memiliki cacat fisik, bisa memanfaatkan *speech recognition* untuk mengendalikan sesuatu yang diinginkannya.

Pengaplikasian ini ingin diterapkan dalam pengendalian kursi roda cerdas ( *smart wheelchair* ). Tujuannya adalah untuk mempermudah seseorang yang memiliki cacat fisik dan memiliki keterbatasan dalam bergerak. Pada Tugas Akhir sebelumnya kondisi dari kursi roda cerdas tersebut telah memiliki kemampuan menghindari rintangan dan sistem kendali menggunakan *joystick*. Dalam Tugas Akhir kali ini akan dibuat pengembangan dari sistem kursi roda cerdas tersebut dengan pengendalian menggunakan ucapan sehingga kursi roda cerdas tersebut dapat lebih bermanfaat bagi seseorang yang memiliki cacat fisik. Proses pengenalan ucapan suara tersebut menggunakan Jaringan Syaraf Tiruan *Multilayer Perceptron* dengan metode pembelajaran *Back Propagation*.

##### 1.2 Tujuan

Tujuan dari penelitian Tugas Akhir ini adalah mengembangkan sistem kendali kursi roda cerdas dengan menggunakan ucapan suara manusia sehingga pemakai kursi roda dengan keterbatasan fisik dapat dengan mudah mengoperasikannya.

##### 1.3 Pembatasan Masalah

Pembatasan masalah pada Tugas Akhir ini adalah sebagai berikut:

1. Ruang gerak kursi roda berada pada daerah yang datar sehingga kursi roda dapat bergerak dengan baik.
2. Mikrokontroler yang digunakan adalah mikrokontroler AVR ATmega128.
3. Bahasa pemrograman yang digunakan adalah bahasa C yang diadaptasikan pada *software Code Vision AVR*.
4. Sensor yang digunakan adalah sensor ultrasonik yang diproduksi oleh Parallax (Sensor PING)))<sup>TM</sup> Ultrasonik Range Finder).
5. Digunakan 3 buah sensor ultrasonik, sehingga terdapat keterbatasan dalam pendeteksian halangan.
6. Pengendalian kecepatan motor DC dilakukan melalui PWM.
7. Suara perintah terdiri dari 5 buah kata bahasa indonesia, yaitu maju, mundur, kanan, kiri dan stop.
8. Data masukan berupa 5 buah kata bahasa Indonesia, yaitu maju, mundur, kanan, kiri, dan stop yang diambil dari 10 orang sebagai sampel suara yang terdiri dari 8 orang pria dan 2 orang wanita. Masing-masing orang diambil datanya sebanyak 10 buah contoh suara untuk setiap kata.

<sup>1</sup> Mahasiswa Jurusan Teknik Elektro UNDIP

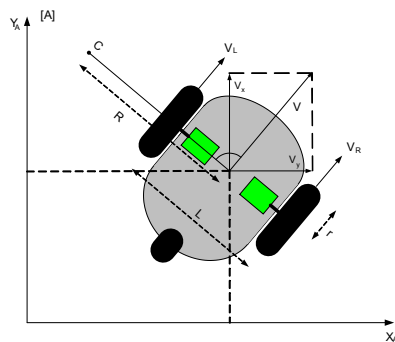
<sup>2</sup> Staf Pengajar Jurusan Teknik Elektro UNDIP

9. Analisis perhitungan sinyal dilakukan dengan ekstraksi ciri menggunakan LPC (*Linear Predictive Coding*) dan FFT (*Fast Fourier Transform*) yang telah tersedia pada *tools* MATLAB.
10. Proses pengenalan suara dilakukan dengan Jaringan Syaraf Tiruan *Multilayer Perceptron* dengan metode pembelajaran *Back Propagation* yang telah tersedia pada *tools* MATLAB.
11. Proses pengiriman data dari komputer melalui komunikasi serial.

## II. DASAR TEORI

### 2.1 Kinematika Kursi Roda Penggerak Differensial

Kursi roda jenis ini pada dasarnya memiliki dua roda utama yang masing-masing digerakan oleh penggerak tersendiri (umumnya berupa motor DC magnet permanen dengan roda gigi reduksi yang berfungsi untuk memperkuat torsi motor), sebagai penyeimbang umumnya kursi roda ini dilengkapi juga dengan satu atau dua buah roda *castor* yang ditempatkan di bagian belakang kursi roda tersebut.



**Gambar 1.** Posisi dan Orientasi Kursi Roda Dalam Sistem Koordinat *Cartesian*.

Secara matematis kecepatan linier dan kecepatan angular kursi roda dirumuskan sesuai persamaan (2.1).

$$\begin{bmatrix} v(t) \\ \omega(t) \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 \\ 1/L & -1/L \end{bmatrix} \begin{bmatrix} v_R(t) \\ v_L(t) \end{bmatrix} \dots\dots (2.1)$$

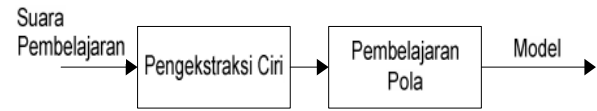
Persamaan (2.1) di atas pada dasarnya memperlihatkan relasi kinematika langsung antara kecepatan linier roda-roda kursi roda terhadap kecepatan linier dan angularnya, sedangkan persamaan (2.9) di bawah memperlihatkan relasi sebaliknya (kinematika balik).

$$\begin{bmatrix} v_R(t) \\ v_L(t) \end{bmatrix} = \begin{bmatrix} 1 & L/2 \\ 1 & -L/2 \end{bmatrix} \begin{bmatrix} v(t) \\ \omega(t) \end{bmatrix} \dots\dots (2.2)$$

### 2.2 Pengenalan Suara

Salah satu bentuk pendekatan pengenalan suara adalah pendekatan pengenalan pola yang terdiri dari dua langkah yaitu pembelajaran pola suara dan pengenalan suara melalui perbandingan pola. Tahap perbandingan pola adalah tahap saat suara yang akan dikenali dibandingkan polanya dengan setiap kemungkinan pola yang telah dipelajari dalam fase pembelajaran, untuk kemudian diklasifikasikan dengan pola terbaik yang

cocok. Blok diagram pembelajaran pola dan pengenalan suara ditunjukkan pada gambar di bawah ini.



(a) Blok diagram pembelajaran pola.



(b) Blok diagram pengenalan suara.

**Gambar 2.** Blok diagram pembelajaran pola dan pengenalan suara.

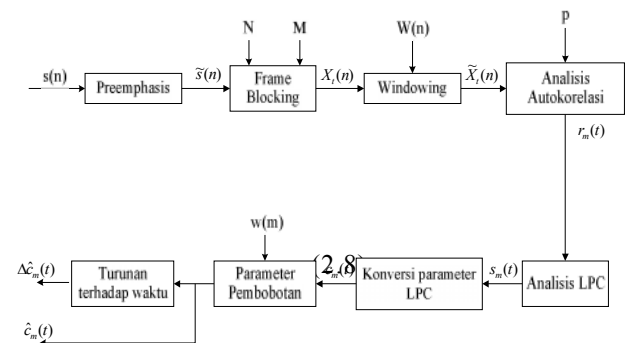
Pengenalan suara secara umum dapat dibagi menjadi tiga tahap, yaitu tahap ekstraksi ciri, tahap pemodelan atau pembelajaran, dan tahap pengenalan suara.

#### 2.2.1 Ekstraksi Ciri dengan Metode LPC

Salah satu metode yang digunakan untuk proses ekstraksi ciri adalah *Linear Predictive Coding* (LPC). Prinsip dasar dari pemodelan sinyal dengan menggunakan LPC adalah bahwa contoh sinyal ucapan  $s(n)$  pada waktu ke- $n$  dapat diperkirakan sebagai kombinasi linier  $p$  sampel sinyal ucapan sebelumnya yaitu

$$s(n) \approx a_1 s(n-1) + a_2 s(n-2) + \dots\dots\dots + a_p s(n-p)$$

dengan koefisien  $a_1, a_2, \dots\dots\dots a_p$  diasumsikan bernilai konstan pada *frame* analisis ucapan. Prosedur untuk mendapatkan koefisien LPC diperlihatkan pada gambar berikut.



**Gambar 3.** Blok diagram analisis LPC.

#### 2.2.2 Algoritma Transformasi Fourier Cepat (*Fast Fourier Transform*)

Untuk lebih memudahkan dalam perhitungan, transformasi Fourier diskret selanjutnya dapat dikembangkan dengan cara mengefisienkan perulangan perhitungan juga sifat simetri dari faktor fasa  $W_N$  yang menjadi sifat transformasi Fourier diskret dengan tujuan mengurangi jumlah perhitungan yang dibutuhkan.

Untuk lebih memudahkan dalam manipulasi atau perhitungan penjumlahan,

$$x(t) = \frac{ao}{2} + \sum_{k=1}^{\infty} (a_k \cos(\omega kt) + b_k \sin(\omega kt))$$

dapat dituliskan kembali dalam bentuk bilangan kompleks eksponensial berdasarkan identitas Euler :

$$\cos(\theta) = \frac{e^{j\theta} + e^{-j\theta}}{2}, \sin(\theta) = \frac{e^{j\theta} - e^{-j\theta}}{2j}$$

Sehingga,

$$x(t) = ao + \sum_{k=1}^{\infty} (a_k \frac{e^{j\omega kt} + e^{-j\omega kt}}{2} + b_k \frac{e^{j\omega kt} - e^{-j\omega kt}}{2j})$$

$$= ao + \sum_{k=1}^{\infty} (\frac{1}{2}(a_k - b_k)e^{j\omega kt} + \frac{1}{2}(a_k + b_k)e^{-j\omega kt})$$

dinyatakan  $y_n = 1/2 (a_n - j b_n)$ , dapat kita gunakan fakta-fakta bahwa  $a_n$  dan  $b_n$  adalah koefisien Fourier fungsi periodik, dengan fungsi kosinus adalah fungsi periodik genap, dan fungsi sinus adalah fungsi periodik ganjil.

Algoritma FFT( *Fast Fourier Transform*) yang banyak digunakan secara luas adalah algoritma dengan bilangan dasar 2, artinya masukan yang dibutuhkan sebanyak  $N = 2^V$ . Hal yang pertama dilakukan dalam pengembangan algoritma FFT (*Fast Fourier Transform*) dengan bilangan dasar 2 ialah dengan menyederhanakan penulisan persamaan DFT (*Discrete Fourier Transform*) dan juga membuat beberapa hubungan matematis, sehingga persamaan dapat ditulis kembali menjadi :

$$X_1(k) = \sum_{n=0}^{N-1} x_n e^{-j2\pi nk/N}, k = 0, 1, \dots, N-1$$

Dan juga faktor  $e^{-j2\pi n/N}$  akan ditulis sebagai  $W_N$ , maka :

$$W_N = e^{-j2\pi n/N}$$

Sehingga persamaan menjadi :

$$X_1(k) = \sum_{n=0}^{N-1} x_n W_N^{kn}, k = 0, 1, \dots, N-1$$

Akan sangat bermanfaat pada saat ini untuk memperhatikan hubungan yang melibatkan  $W_N$  pertama

$$W_N^2 = (e^{-j2\pi/N})^2 = e^{-j2\pi 2/N} = e^{-j2\pi(N/2)} = W_{N/2}$$

Kedua

$$W_N^{(K+N/2)} = W_N^K W_N^{N/2} = W_N^K e^{-j(2\pi/N)(N/2)}$$

$$= W_N^K e^{-j\pi} (\cos \pi - j \sin \pi)$$

$$= -W_N^K$$

bila hasilnya disimpulkan yang berhubungan dengan  $W_N$  maka kita dapatkan

$$W_N = e^{-j2\pi/N}$$

$$W_N^2 = W_{N/2}$$

$$W_N^{(K+N/2)} = -W_N^K$$

Prosedur ini dijelaskan sebagai berikut : sufiks n, dalam persamaan (2.28) terletak dari  $n = 0$  hingga  $n = N-1$ , berhubungan dengan nilai data  $X_0, X_1, X_2, X_3, \dots, X_{N-1}$ . Runtun nomor genap adalah  $X_0, X_2, X_4, X_6, \dots, X_{N-2}$  dan runtun bernomor ganjil adalah  $X_1, X_3, X_5, \dots, X_{N-1}$ , kedua runtun mengandung  $N/2$  poin. Istilah dalam runtun genap dapat diwakili dengan  $X_{2n}$  dengan  $n = 0$  hingga  $n = N/2 - 1$  sementara untuk runtun ganjilnya menjadi  $X_{2n+1}$ , maka persamaan  $X_1(k)$  dapat ditulis ulang menjadi :

$$x_1(k) = \sum_{n=0}^{\frac{N}{2}-1} x_{2n} W_N^{2nK} + \sum_{n=0}^{\frac{N}{2}-1} x_{2n+1} W_N^{(2n+1)k}$$

$$= \sum_{n=0}^{\frac{N}{2}-1} x_{2n} W_N^{2nK} + W_N^K \sum_{n=0}^{\frac{N}{2}-1} x_{2n+1} W_N^{2nK}, k = 0, 1, \dots, N-1$$

dengan menggunakan persamaan (2.31b) memberikan

$W_N^{2nK} = W_{N/2}^{nk}$  sehingga persamaan (2.32) menjadi :

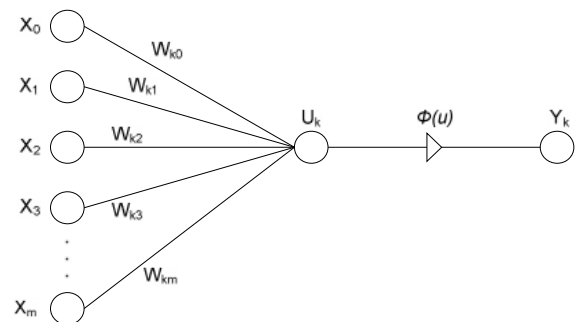
$$x(k) = \sum_{n=0}^{\frac{N}{2}-1} x_{2n} W_{N/2}^{nk} + W_N^K \sum_{n=0}^{\frac{N}{2}-1} x_{2n+1} W_{N/2}^{nk}, k = 0, 1, \dots, N-1$$

persamaan dapat ditulis menjadi :

$$X_1(k) = X_{11}(k) + W_N^K X_{12}(k), k = 0, 1, \dots, N-1$$

### 2.2.3 Pemodelan Jaringan Syaraf Tiruan

Model matematis dari neuron dapat dilihat pada Gambar 4. Masukan dan bobot merupakan model matematis dari dendrit, penjumlahan serta fungsi pengaktif merupakan model matematis dari soma (badan sel), serta keluaran merupakan model matematis dari akson.



Gambar 4. Model matematis neuron

Setiap jalur nilai masukan ke neuron mempunyai bobot, dan terdapat satu jalur masukan tambahan yang bukan merupakan jalur masukan, yaitu jalur ke-0. Jalur ke-0 ini dapat berfungsi sebagai nilai koreksi atau dapat juga berfungsi sebagai nilai *threshold* (nilai ambang batas). Impuls pada sistem syaraf manusia dimodelkan sebagai nilai yang dimasukkan melalui jalur nilai masukan ke dalam model neuron. Nilai pada setiap jalur masukan akan dikalikan dengan bobot jalur, dan kemudian dijumlahkan. Nilai yang dihasilkan dari penjumlahan seluruh masukan tersebut dapat berada pada *range* yang sangat besar sehingga akan mempersulit pengolahan selanjutnya. Untuk menghindari hal tersebut, nilai hasil penjumlahan akan diproses terlebih dulu dengan menggunakan sebuah fungsi  $\phi$  (phi) yang disebut fungsi aktivasi. Penghitungan nilai hasil dapat dituliskan secara matematis sebagai berikut:

$$y_k = \phi(u)$$

$$y_k = \phi(\sum_{i=0}^m x_i w_{ki})$$

Proses pembelajaran pada suatu neuron dilakukan dengan memperbarui nilai bobot setiap masukan ( $w_i$ )

sesuai dengan pembelajaran yang diberikan. Bentuk data pembelajaran adalah berupa pasangan vektor nilai masukan ke neuron dan nilai target keluaran ( $t$ ). Nilai masukan akan dimasukkan ke dalam neuron, dan menghasilkan nilai keluaran ( $y$ ). Kesalahan dari nilai keluaran yang dihasilkan dihitung, dan nilai bobot  $w$  untuk setiap masukan diperbarui menggunakan:

$$w_i(n+1) = w_i(n) + \Delta w_i$$

$$\Delta w_i = \eta(t - y)x_i$$

Dengan  $\eta(\mu)$  adalah kecepatan pembelajaran, yaitu seberapa besar bobot  $w$  terbaru pada setiap langkah pembelajaran. Proses ini dilakukan berulang-ulang dengan data pembelajaran yang berbeda. Pada akhirnya, bobot nilai masukan akan berisi nilai bobot yang sesuai.

#### 2.2.4 Metode Pelatihan Backpropagation

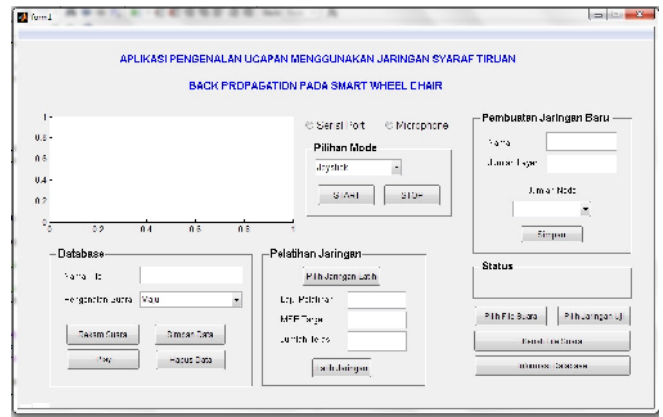
Algoritma pelatihan *backpropagation* meliputi 3 fase. Fase pertama adalah fase maju. Pola masukan dihitung maju mulai dari *layer* masukan hingga *layer* keluaran menggunakan fungsi aktivasi yang ditentukan. Fase kedua adalah fase mundur. Selisih antara keluaran jaringan dengan target yang diinginkan merupakan kesalahan yang terjadi. Kesalahan tersebut dipropagasikan mundur, dimulai dari garis yang berhubungan langsung dengan unit-unit di *layer* keluaran. Fase ketiga adalah modifikasi bobot untuk menurunkan kesalahan yang terjadi.

Ketiga fase tersebut diulang-ulang terus hingga kondisi penghentian dipenuhi. Umumnya kondisi penghentian yang sering dipakai adalah jumlah iterasi atau kesalahan. Iterasi akan dihentikan jika jumlah iterasi yang dilakukan sudah melebihi jumlah maksimum iterasi yang ditetapkan, atau jika kesalahan yang terjadi sudah lebih kecil dari batas toleransi kesalahan yang ditetapkan.

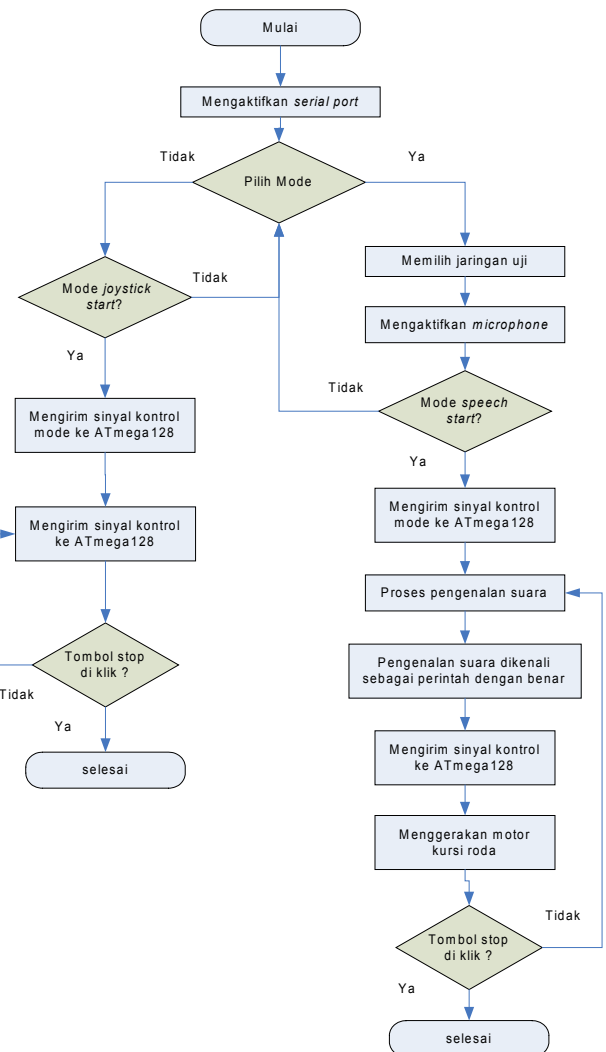
### III. PERANCANGAN SISTEM

#### 3.1 Perancangan Perangkat Lunak

Perancangan perangkat lunak menggunakan bahasa pemrograman Matlab 2009a. Perangkat lunak ini berfungsi untuk menentukan mode yang akan digunakan untuk mengendalikan kursi roda yaitu mode *joystick* dan mode *speech recognition*. Perangkat lunak pada perancangan sistem ini juga berfungsi untuk pelatihan jaringan syaraf tiruan serta melakukan proses pengolahan suara sehingga dihasilkan pengenalan ucapan yang sesuai dan digunakan sebagai parameter pada mode *speech* yang akan dikirim sebagai parameter perintah ke mikrokontroler.



Gambar 5. Tampilan visual program utama kursi roda cerdas.



Gambar 6. Diagram alir program utama kursi roda cerdas.

##### 3.1.1 Perancangan Mode Speech Recognition

Pada perancangan mode *speech*, mode ini akan aktif apabila telah memilih jaringan uji, mengaktifkan *microphone*, memilih mode *speech* pada pop-up menu dan tombol *start* ditekan. Akan mati apabila tombol *stop* ditekan. Selain mengirimkan sinyal kontrol untuk mengaktifkan fungsi *speech*, *Notebook* juga mengirimkan sinyal kontrol posisi hasil pengolahan dan pengenalan suara ke mikrokontroler sebagai acuan gerakan motor kursi roda.

Ketika tombol *start* ditekan, *Notebook* akan mengirimkan data karakter 50 (desimal) ke mikrontroler.

Pada perancangan mode *speech* ini, sebelum didapatkan sebuah pengenalan ucapan yang tepat dan benar maka perlu adanya proses pengelolaan basis data suara, proses pengolahan suara, dan pelatihan jaringan syaraf tiruan yang tepat supaya didapatkan pengenalan ucapan dengan tingkat keberhasilan pengenalan yang baik. Sebelum membahas pada sub bab selanjutnya, berikut adalah diagram alur tahapan proses pembuatan jaringan syaraf tiruan yang akan dijadikan sebagai jaringan uji pada mode *speech*.



**Gambar 7.** Diagram alur tahapan proses pembuatan jaringan syaraf tiruan.

### 3.1.2 Pengelolaan Basis Data

Dalam pembuatan basis data dilakukan perekaman suara dengan 10 orang responden yang terdiri dari 8 orang pria dan 2 orang wanita dengan rentang usia yang bervariasi. Tiap responden melakukan perekaman sebanyak 10 kali untuk setiap kata, sehingga diperoleh 50 suara basis data untuk tiap responden. *Database* ditambah dengan perekaman suara di luar kata "maju", "mundur", "kanan", "kiri", dan "stop" yang akan dikenali sebagai "Suara Tidak Dikenali" sebanyak 50 suara, sehingga jumlah total keseluruhan basis data ialah 550 suara. Perekaman suara dalam pembuatan basis data ini dilakukan selama 1 detik untuk tiap ucapan.

Pencuplikan dilakukan pada frekuensi 8000 Hz pada resolusi 16 bit. Kecepatan pencuplikan tersebut dilakukan dengan didasarkan asumsi bahwa sinyal percakapan (*speech*) berada pada daerah frekuensi 300-3400 Hz sehingga memenuhi kriteria Nyquist yang menyatakan bahwa frekuensi penyampling minimal 2 kali frekuensi maksimal dari sinyal yang akan disampling untuk menghindari *aliasing* :

$$f_s \geq 2 \times f_h \quad f_h = f_{in \text{ tertinggi}}$$

Sedangkan untuk data uji yaitu data yang digunakan untuk proses pengujian jaringan, diambil dari 5 orang responden, 4 pria dan 1 wanita. Data uji diambil dari 2 orang di dalam *database* dan 3 orang di luar *database* sehingga total ada 250 data yang diujikan. Selain itu juga akan dilakukan pengujian secara online terhadap unjuk kerja jaringan nantinya.

### 3.1.3 Ekstraksi Ciri

Proses pengekstraksian ciri merupakan proses untuk mendapatkan parameter-parameter sinyal suara. Parameter-parameter inilah yang nantinya digunakan untuk membedakan satu kata dengan kata yang lain. Ada beberapa cara dalam memperoleh ciri suara dan dalam Tugas Akhir ini proses ekstraksi ciri dilakukan dengan metode LPC (*Linear Prediction Coding*).

Proses awal ekstraksi ciri yaitu dengan melakukan normalisasi amplitudo terlebih dahulu untuk mengatasi jarak antara mulut dengan *microphone*, setelah itu dilakukan pendeteksian awal dan akhir kata. Setelah itu sinyal disamakan panjang datanya dengan normalisasi panjang data.

Normalisasi amplitudo dilakukan dengan cara membagi semua nilai input dengan nilai maksimum dari input itu sendiri. Sehingga untuk semua sinyal masukan memiliki nilai maksimum yang sama yaitu 1. Hal ini digunakan untuk mengatasi jarak dekat atau jauhnya mulut dengan *microphone*.

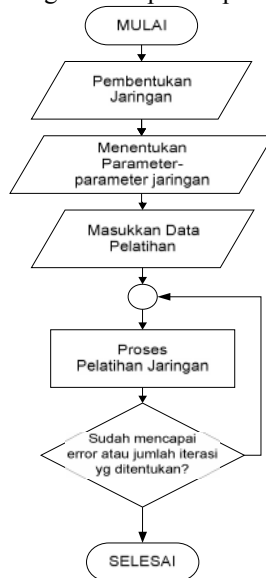
Sedangkan deteksi awal-akhir digunakan pada proses untuk mendeteksi mulai sinyal ucapan awal dan berakhir ketika sudah tidak diucapkan. Sehingga tidak disalah artikan sebagai sinyal yang masuk. Setelah itu sinyal terlebih dahulu disamakan jumlah datanya, sehingga pada proses analisa LPC bisa didapat jumlah frame dan panjang masing-masing frame yang sama. Hal ini akan menyebabkan proses pengenalan suara semakin penjang namun dapat memperkecil kesalahan pengenalan nantinya.

Proses selanjutnya yaitu dengan menganalisa suara dengan metode LPC sehingga diperoleh beberapa koefisien LPC yang merupakan ciri dari suatu suara. Selanjutnya koefisien LPC yang diperoleh tersebut diproses dengan *Fast Fourier Transform* (FFT) untuk mendapatkan sinyal pada domain frekuensi. Hal ini bertujuan agar perbedaan antar pola kata yang satu dengan yang lain terlihat lebih jelas sehingga ekstraksi parameter sinyal memberikan hasil yang lebih baik. Hasil keluaran dari FFT ini akan dijadikan masukan bagi jaringan syaraf tiruan *Back Propagation* di mana jaringan syaraf tiruan ini berfungsi sebagai bagian utama dari sistem untuk proses pengenalan suara.

Proses *Fast Fourier Transform* (FFT) yang digunakan memakai 512 *point* dan karena hasil FFT simetris maka keluaran FFT tersebut hanya diambil setengah saja yang dianggap sudah dapat merepresentasikan ciri dari sinyal ucapan dalam domain frekuensi. Dari data FFT ini kemudian dijadikan masukan bagi jaringan syaraf tiruan.

### 3.1.4 Pelatihan Jaringan

Setelah melalui tahap transformasi Fourier selanjutnya parameter-parameter yang dihasilkan dimasukkan ke dalam jaringan syaraf tiruan dengan menggunakan metode pembelajaran *Back Propagation*. Proses pelatihan jaringan syaraf tiruan dilakukan dengan mengambil input hasil proses FFT dari seluruh suara hasil perekaman pada basis data. Jumlah suara pada basis data adalah 550 kata, sehingga input pelatihan bagi JST sebanyak data FFT x 550 data. Berikut adalah diagram alir proses pelatihan jaringan.



Gambar 8. Diagram alir proses pelatihan jaringan

Selain menentukan masukan pelatihan kita juga harus menentukan *output target* pelatihan. Dalam sistem ini, fungsi aktivasi yang digunakan ialah *sigmoid bipolar* yang memiliki range keluaran antara -1 sampai 1, yaitu sebagai berikut :

- maju = [-1; 1; 1; 1; 1; 1; 1; 1; 1; 1]
- mundur = [1; 1;-1; 1; 1; 1; 1; 1; 1; 1]
- kanan = [1; 1; 1; 1;-1; 1; 1; 1; 1; 1]
- kiri = [1; 1; 1; 1; 1; 1;-1; 1; 1; 1]
- stop = [1; 1; 1; 1; 1; 1; 1; 1;-1; 1]
- tidak dikenali = [-1;-1;-1; 1; 1; 1;-1; 1;-1; 1]

### 3.1.5 Pengenalan Ucapan

Proses pengenalan ucapan diawali dengan memasukkan perintah suara melalui *microphone*, kemudian dilakukan proses normalisasi pada suara tersebut. Selanjutnya terjadi proses ekstraksi ciri dan transformasi Fourier dari suara uji hingga diperoleh data akhir yang akan diujikan pada jaringan yang telah ditentukan. Jaringan akan memutuskan suara tersebut dikenali sebagai kata maju, mundur, kanan, kiri atau stop. Selanjutnya hasil keputusan jaringan tersebut diubah dalam bentuk karakter yang akan dikirimkan pada kursi roda. Kata maju akan diubah dalam karakter '5', mundur dalam karakter '4', kanan dalam karakter '3', kiri dalam karakter '2', dan stop dalam karakter '1'. Data ini kemudian dikirimkan pada kursi roda melalui komunikasi serial.

Sebelum mengirimkan data melalui komunikasi serial, kita perlu membuat objek serial pada program Matlab dengan menggunakan perintah sebagai berikut :

```

s1 = serial
('COM1','baudrate',9600,'DataBits',8);
s1.StopBits = 1;
s1.Parity = 'none';

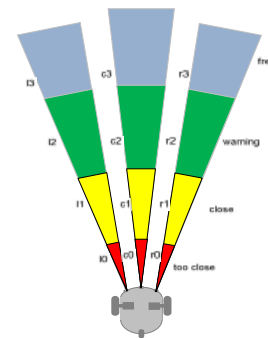
```

Sebelum mengirimkan data, objek serial ini harus dibuka terlebih dahulu dengan perintah `fopen(s1)`. Setelah itu data kita kirimkan dengan perintah `fwrite(s1,5,'char')` misalnya untuk memerintahkan kursi roda bergerak maju. Untuk menutup port serial digunakan perintah `fclose(s1)`.

### 3.2 Obstacle Avoidance

Merupakan subrutin program yang berisi perintah untuk mengatur kursi roda agar bergerak menghindari halangan yang berada di depannya, dengan cara berbelok ke kiri atau kanan. Perancangan tingkah laku *obstacle avoidance* dilakukan sesuai berdasar pada pembagian *range* pada tiap sensor yaitu *too close*, *close*, *warning* dan *free*.

Daerah *too close* adalah daerah yang berada pada jarak kurang dari 45 cm. daerah *close* adalah daerah yang berada pada jarak diantara 45 cm sampai 100 cm. daerah *warning* adalah daerah yang berada pada jarak diantara 100 cm sampai 150 cm. Sedangkan daerah *free* adalah daerah yang berada pada jarak lebih dari 150 cm. Gambar 10 merupakan gambar pembagian area dari sensor ping.



Gambar 9. Gambar pembagian area dari sensor ping.

Tabel 1 Posisi halangan pada setiap keadaan

Keadaan	Posisi halangan
Keadaan 1	
Keadaan 2	
Keadaan 3	
Keadaan 4	
Keadaan 5	

## IV. PENGUJIAN DAN ANALISA

### 4.1 Pengujian Perangkat Lunak

#### 4.1.1 Analisa Pelatihan Jaringan

Penentuan jaringan terbaik ini dapat ditunjukkan melalui besarnya MSE (*Mean Square Error*). Di mana semakin kecil MSE maka tingkat keberhasilan pengenalan akan semakin besar. Berikut adalah Tabel 4 variasi jaringan pada saat pelatihan jaringan.

**Tabel 2** Variasi jaringan yang digunakan dalam pelatihan

Nama	Struktur Jaringan	lr	MSE Target	Jumlah Iterasi
tes_fix 1	3 hidden layer (300, 150, 50)	0,05	$10^{-50}$	10.000
tes_fix 2	4 hidden layer (350, 225, 125, 50)	0,05	$10^{-50}$	10.000
tes_fix 3	5 hidden layer (350, 250, 150, 100, 50)	0,05	$10^{-50}$	10.000

Perbandingan hasil pelatihan ketiga jaringan di atas dapat kita lihat pada tabel berikut.

**Tabel 3** Perbandingan hasil pelatihan

Nama	Struktur Jaringan	lr	Jumlah iterasi	MSE akhir
tes_fix 1	3 hidden layer (300, 150, 50)	0,05	1173	$8,52 e^{-33}$
tes_fix 2	4 hidden layer (350, 225, 125, 50)	0,05	863	$2,47 e^{-33}$
tes_fix 3	5 hidden layer (350, 250, 150, 100, 50)	0,05	743	$2,38 e^{-33}$

Dengan melihat data perbandingan hasil pelatihan di atas, dapat dilihat bahwa dari variasi jumlah *neuron* dan *hidden layer* pada struktur jaringan diperoleh kesimpulan bahwa semakin kompleks jaringan yang ditandai dengan semakin banyak lapisan suatu jaringan, maka MSE terkecil akan dicapai dengan jumlah iterasi pelatihan yang semakin sedikit. Namun waktu yang diperlukan untuk proses pelatihan juga akan semakin lama. Karena semakin rumit perhitungan pada saat pelatihan jaringan.

#### 4.1.2 Analisa Tahap Pengenalan

Pada pengujian ini digunakan 5 orang responden, 3 orang dari luar database, dan 2 orang dari dalam database. Selanjutnya dari kelima responden itu direkam masing-masing 10 pengucapan untuk tiap kata, sehingga total ada 250 data yang diujikan. Dari hasil pengujian diperoleh hasil sebagai berikut :

**Tabel 4** Hasil pengujian ketiga jaringan dengan data uji responden di dalam database

Jaringan	Persentase Pengenalan %
tes_fix1	84 %
tes_fix2	86 %
tes_fix3	87 %

**Tabel 5** Hasil pengujian ketiga jaringan dengan data uji responden di luar database

Jaringan	Persentase Pengenalan %
tes_fix1	83,33 %
tes_fix2	82 %
tes_fix3	85,33 %

Dari 2 tabel di atas dapat kita lihat bahwa prosentase pengenalan dengan responden di luar database lebih kecil daripada prosentase pengenalan dengan responden di dalam database. Hal ini disebabkan oleh adanya beberapa faktor, diantaranya faktor derau, karakteristik sinyal suara yang dikenali, dan adanya pengaruh distorsi

#### 4.1.3 Pengujian Tingkah Laku Kursi Roda Tanpa Halangan

Setelah melakukan pengujian pada sistem Pada pengujian ini, ketiga sensor PING)))<sup>TM</sup> tidak mendeteksi adanya halangan, dan kursi roda dikendalikan dengan menggunakan perintah ucapan. Berikut adalah pengujian yang dilakukan dengan memberikan perintah ucapan ke kursi roda sebanyak 20 kali secara acak dengan perintah yang telah di tentukan. Hasil pengujian dapat dilihat pada Tabel 8 berikut ini :

**Tabel 6** Hasil pengujian dengan respon kursi roda

No.	Perintah	Gerak Kursi Roda
1	Maju	Maju
2	Stop	Stop
3	Mundur	Mundur
4	Stop	Stop
5	Kanan	Kanan
6	Kiri	Kiri
7	Stop	Stop
8	Maju	Maju
9	Kiri	Kiri
10	Mundur	Mundur
11	Kanan	Kanan
12	Kiri	Kiri
13	Mundur	Mundur
14	Stop	Stop
15	Maju	Maju
16	Kanan	Kanan
17	Kiri	Kiri
18	Mundur	Mundur
19	Kiri	Kiri
20	Stop	Stop

#### 4.1.4 Pengujian Tingkah Laku Kursi Roda Dengan Halangan

Pada tahap ini, pengujian dilakukan dengan meletakkan benda di beberapa posisi yang dianggap sebagai halangan untuk mengindikasikan keadaan kursi roda terhadap halangan. Kursi roda ini mempunyai 5 aturan yang mewakili dari 5 keadaan yang berbeda-beda. Pada pengujian kali ini hanya dilakukan 3 keadaan saja. Kelima aturan ini hanya aktif pada saat kursi roda digerakkan maju.

##### 1. Keadaan 1

Keadaan ini terjadi apabila ketiga sensor, kedua sensor, atau salah satu sensor mendeteksi adanya objek pada jarak kurang dari 45 cm. Pada keadaan ini kursi roda hanya bisa digerakkan ke belakang (mundur).





**Gambar 10.** Hasil pengujian tingkah laku kursi roda pada keadaan 1

## 2. Keadaan 3

Keadaan ini terjadi apabila sensor kanan atau kiri, sensor tengah dan kanan, atau sensor tengah dan kiri mendeteksi adanya objek pada jarak diantara 45 cm sampai 100 cm. Kursi roda akan berbelok secara otomatis ke arah di mana sensor ping mendeteksi jarak yang lebih jauh dengan kecepatan yang semakin lambat.



**Gambar 11.** Hasil pengujian tingkah laku kursi roda pada keadaan 3

## 3. Keadaan 4

Keadaan ini terjadi apabila sensor tengah kursi roda mendeteksi adanya halangan pada jarak antara 100 cm sampai 150 cm. Pada keadaan ini kursi roda akan bergerak maju dengan kecepatan yang berkurang seiring jarak halangan yang semakin dekat.



**Gambar 12.** Hasil pengujian tingkah laku kursi roda pada keadaan 4

## V. PENUTUP

### 5.1 Kesimpulan

Dari hasil penelitian dan pembahasan dapat disimpulkan bahwa :

1. Nilai persentase pengenalan masing-masing jaringan yang diperoleh dengan data uji dari responden di luar database adalah tes\_fix1 = 83,33%, tes\_fix2 = 82%, tes\_fix3 = 85,33%. Sedangkan nilai persentase pengenalan masing-masing jaringan yang diperoleh dengan data uji dari responden dari dalam database adalah tes\_fix1 = 84%, tes\_fix2 = 86%, tes\_fix3 = 87%.
2. Proses perekaman suara sampel harus dilakukan dengan ucapan pelafalan yang tepat dan tidak dibuat-buat karena akan berpengaruh pada hasil pengenalan ucapan.
3. Proses perekaman suara sampel yang bersih dari derau akan membuat tingkat pengenalan ucapan yang baik.
4. Keberhasilan pengenalan ucapan juga dipengaruhi oleh microphone yang digunakan. Microphone dengan fitur pengurangan distorsi akan membuat pengenalan ucapan yang baik.
5. Kesalahan pengenalan yang terjadi diakibatkan adanya perbedaan yang terlalu besar antara sinyal suara yang hendak dikenali dengan sinyal suara yang dilatihkan, hal ini dapat diatasi dengan

menambahkan/memperbanyak berbagai variasi pola kata pada saat pelatihan dengan demikian sistem jaringan lebih diperkaya pengetahuannya.

6. Pada penelitian ini sistem jaringan syaraf tiruan yang dibuat masih terdapat kelemahan yaitu ketika suara masukan diluar kelima kata perintah masih terdeteksi sebagai masukan kata perintah.
7. Penambahan suara sampel data pelatihan jaringan akan memperbaiki kinerja jaringan syaraf tiruan dalam pengenalan ucapan.

### 5.2 Saran

Adapun saran yang dapat diberikan sehubungan dengan pelaksanaan penelitian ini adalah :

1. Karena sistem sangat peka terhadap sinyal derau, maka agar dapat dipertahankan keberhasilan pengenalnya, diperlukan suatu ruangan yang cukup tenang dan bersih terhadap sinyal derau.
2. Perekaman dilakukan dengan suara responden yang asli (menggunakan pelafalan yang baik dan benar serta intonasi yang tepat), tidak dalam keadaan yang dibuat-buat.
3. Untuk tahap selanjutnya bisa dikembangkan sistem modul pengenalan ucapan secara tertanam pada kursi roda, sehingga tidak diperlukan lagi sebuah *Notebook* sebagai alat pemroses pengenalan ucapan.
4. Penambahan sensor jarak pada kursi roda cerdas ini akan memperluas area pendeteksian halangan.
5. Penambahan sensor *limit switch* pada sekeliling kursi roda dapat mendeteksi letak terjadinya benturan/tabrakan.

## DAFTAR PUSTAKA

- [1] Adipranata, Rudy dan Resmana, *Pengenalan Suara Manusa dengan Metode LPC dan Jaringan Syaraf Tiruan Propagasi Balik*, Universitas Kristen Petra, Surabaya, 1999.
- [2] Drs. Jong Jek Siang, M. Sc., *Jaringan Syaraf Tiruan & Pemrogramannya Menggunakan Matlab*, ANDI, Yogyakarta, 2004.
- [3] Setiawan, Iwan, *Fungsi Soft Timer untuk Keperluan Operasi Tundaan dan Penjadwalan (Scheduling) Pada Sistem Embedded*, <http://iwan.blog.undip.ac.id>, Januari 2010.
- [4] Miftahurrozaq, *Perancangan Sistem Kursi Roda Cerdas Menggunakan Pendekatan Model Fungsional dan Model Tingkah Laku dan Perancangan Perilaku Obstacle Avoidance*, Skripsi S-1, Universitas Diponegoro, Semarang, 2010.
- [5] Irfandy, Mahmud, *Aplikasi Pengenalan Ucapan Dengan Jaringan Syaraf Tiruan Propagasi Balik Untuk Pengendalian Robot Bergerak*, Skripsi S-1, Universitas Diponegoro, Semarang, 2009.
- [6] Muhammad, Yajid, *Analisis Pengenalan Ucapan Dengan Ekstraksi MFCC Melalui Jaringan Syaraf Tiruan LVQ*, Skripsi S-1, Universitas Diponegoro, Semarang, 2005.
- [7] -----, (PING)))<sup>TM</sup> *Ultrasonic Distance Sensor Data Sheet*, <http://www.parallax.com>.
- [8] -----, *Using Analog Input Trigger*, <http://www.mathwork.com>.
- [9] -----, *Atmega 128 Data Sheet*, <http://www.atmel.com>.



**Biodata Penulis**

Penulis lahir di kota Semarang pada 16 April 1987. Sekarang tengah menyelesaikan pendidikan Strata Satu di konsentrasi Teknik Kontrol, Jurusan Teknik Elektro, Fakultas Teknik, Universitas Diponegoro, Semarang.

Menyetujui,

Dosen Pembimbing I,

Dosen Pembimbing II,

Achmad Hidayatno, ST, MT  
**NIP.196912211995121001**

Iwan Setiawan, ST, MT  
**NIP.197309262000121001**

Tanggal : \_\_\_\_\_

Tanggal : \_\_\_\_\_